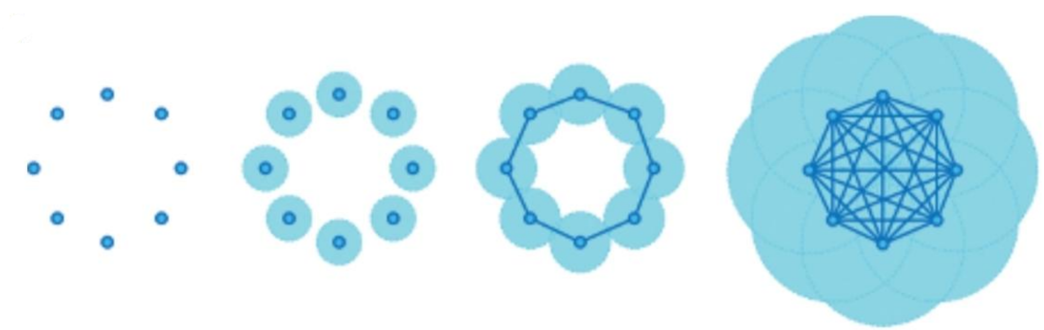
Week 10

Video: https://youtu.be/OkDs9Wj5G1U

Assume that we are given experimental data in the form of a finite metric space S; there are points or vectors that represent measurements along with some distance function (e.g., given by a correlation or a measure of dissimilarity) on the set of points or vectors. Whether or not the set S is a sample from some underlying topological space, it is useful to think of it in those terms. Our goal is to recover the properties of such an underlying space in a way that is robust to small perturbations in the data S. In a broad sense, this is the subject of *topological inference*. (See [82] for an overview.) If S is a subset of Euclidean space, one can consider a 'thickening' S_ϵ of S given by the union of balls of a certain fixed radius ϵ around its points and then compute the Čech complex. One can thus try to compute qualitative features of the data set S by constructing the Čech complex for a chosen value ϵ and then computing its simplicial homology. The problem with this approach is that there is a priori no clear choice for the value of the parameter ϵ. The key insight of PH is the following: To extract qualitative information from data, one considers several (or even all) possible values of the parameter ϵ. As the value of ϵ increases, simplices are added to the complexes. Persistent homology then captures how the homology of the complexes changes as the parameter value increases, and it detects which features 'persist' across changes in the parameter value.

(S, d)  $\subseteq \mathbb{R}^2$



A roadmap for the computation of PH, Otter et. al.

Def: Let K be a r-simplicial complex. A filtration $\mathcal{F}$ of K is a finite sequence of nested subcomplexes of K (not its skeleta necessarily):

$$\mathcal{F}: K_0 \subseteq K_1 \subseteq \cdots \subseteq K_n \subseteq \cdots \subseteq K_r = K.$$

Ex:

1. The skeleta of K is a filtration of K.

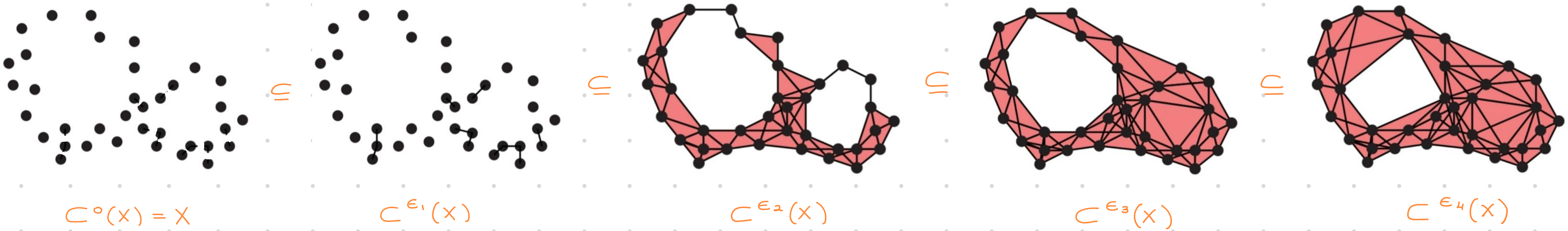
2. For $0 < \epsilon_1 < \epsilon_2 < \cdots < \epsilon_r$ and (X, d) a finite metric space in $\mathbb{R}^N$

$$\mathcal{C}: C^0(X) \subseteq C^{\epsilon_1}(X) \subseteq C^{\epsilon_2}(X) \subseteq \cdots \subseteq C^{\epsilon_r}(X)$$

$$\mathcal{V}: VR^0(X) \subseteq VR^{\epsilon_1}(X) \subseteq VR^{\epsilon_2}(X) \subseteq \cdots \subseteq VR^{\epsilon_r}(X)$$

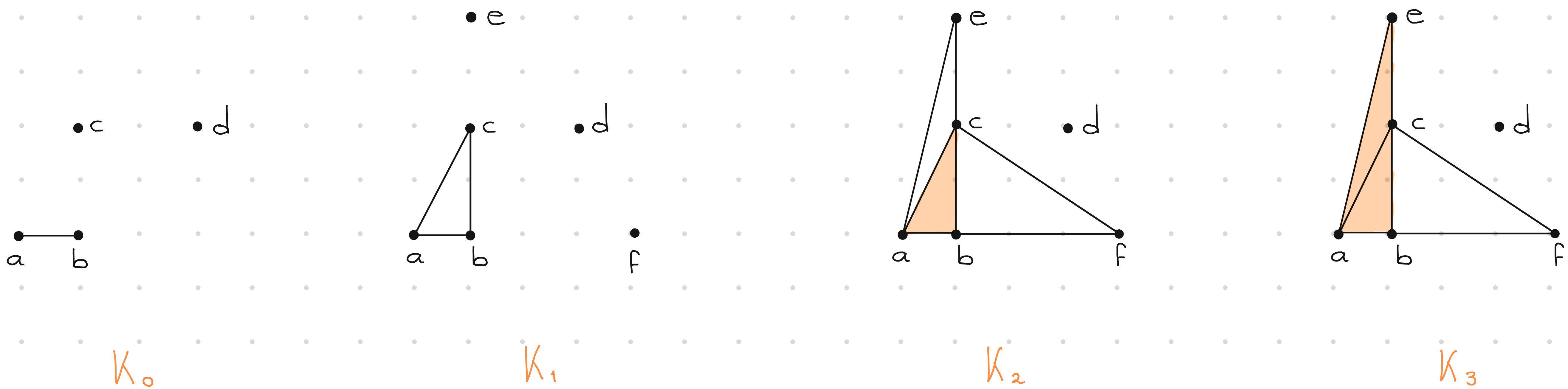
are filtrations of $C^{\epsilon_r}(X)$ and $VR^{\epsilon_r}(X)$, respectively.

[1]



$C^0(X) = X$
$\subseteq$
$C^{\epsilon_1}(X)$
$\subseteq$
$C^{\epsilon_2}(X)$
$\subseteq$
$C^{\epsilon_3}(X)$
$\subseteq$
$C^{\epsilon_4}(X)$

3.



a b c d
K_0
a b c d e f
K_1
a b c d e f
K_2
a b c d e f
K_3

Now, for each subcomplex K_i in the filtration, we can take its homology for any degree p: $H_0(K_i;\mathbb{F}_2)$, $H_1(K_i;\mathbb{F}_2)$, ..., $H_p(K_i;\mathbb{F}_2)$, ...

Moreover, for all $p \geq 0$ the inclusion $K_i \hookrightarrow K_j$ induces a linear transformation

$$f_{ij} : H_p(K_i;\mathbb{F}_2) \longrightarrow H_p(K_j;\mathbb{F}_2) \quad \text{for all } i,j \in \{0,1,\dots,r\}.$$

$$\mathcal{F}: \quad K_0 \hookrightarrow K_1 \hookrightarrow \cdots \hookrightarrow K_r$$

$$H_*(\mathcal{F}): \quad \begin{array}{ccccccc}
H_0(K_0;\mathbb{F}_2) & \xrightarrow{f_{01}} & H_0(K_1;\mathbb{F}_2) & \xrightarrow{f_{12}} & \cdots & \xrightarrow{f_{r-1,r}} & H_0(K_r;\mathbb{F}_2) \\
H_1(K_0;\mathbb{F}_2) & \xrightarrow{f_{01}} & H_1(K_1;\mathbb{F}_2) & \xrightarrow{f_{12}} & \cdots & \xrightarrow{f_{r-1,r}} & H_1(K_r;\mathbb{F}_2) \\
H_p(K_0;\mathbb{F}_2) & \xrightarrow{f_{01}} & H_p(K_1;\mathbb{F}_2) & \xrightarrow{f_{12}} & \cdots & \xrightarrow{f_{r-1,r}} & H_p(K_r;\mathbb{F}_2)
\end{array}$$

⚠ Observe that i, j do not need to be consecutive. We could also have

$$K_1 \hookrightarrow K_{10} \quad \text{and} \quad f_{1,10} : H_p(K_1; \mathbb{F}_2) \longrightarrow H_p(K_{10}; \mathbb{F}_2).$$

From functoriality, we have the following "compatibility" property: Why?

$$\begin{tikzcd} K_a \arrow[rr, hook] \arrow[dr, hook] & & K_c \\ & K_b \arrow[ur, hook] & \end{tikzcd} \implies \begin{tikzcd} H_p(K_a; \mathbb{F}_2) \arrow[rr, "f_{ac}"] \arrow[dr, "f_{ab}"'] & \circlearrowleft & H_p(K_c; \mathbb{F}_2) \\ & H_p(K_b; \mathbb{F}_2) \arrow[ur, "f_{bc}"'] & \end{tikzcd}$$

$$f_{bc} \circ f_{ab} = f_{ac}$$

Def: Let $\mathcal{F}: K_0 \subseteq K_1 \subseteq \cdots \subseteq K_r = K$ be a filtration for K. The *p-th homology of $\mathcal{F}$* is the pair

$$H_p(\mathcal{F}) : \left(\{ H_p(K_i; \mathbb{F}_2) \}_{i=0}^{r} ; \{ f_{ij} \}_{0 \leq i \leq j \leq r} \right),$$

where for all $0 \le i \le j \le r$ the linear transformations $f_{ij}: H_p(K_i;\mathbb{F}_2) \longrightarrow H_p(K_j;\mathbb{F}_2)$

are induced by the inclusion $K_i \hookrightarrow K_j$.

$$H_p(\mathcal{F}): \quad H_p(K_0) \longrightarrow H_p(K_1) \longrightarrow \cdots \longrightarrow H_p(K_i) \xrightarrow{f_{ij}} \cdots \longrightarrow H_p(K_j) \longrightarrow \cdots \longrightarrow H_p(K).$$

Def: Let $\mathcal{F}: K_0 \subseteq K_1 \subseteq \cdots \subseteq K_r = K$ be a filtration for K.

① The **(i,j)-persistent p-th homology group of K** is the image of f_{ij} in the p-th homology of $\mathcal{F}$, and it is denoted by $H_p^{i,j}$:

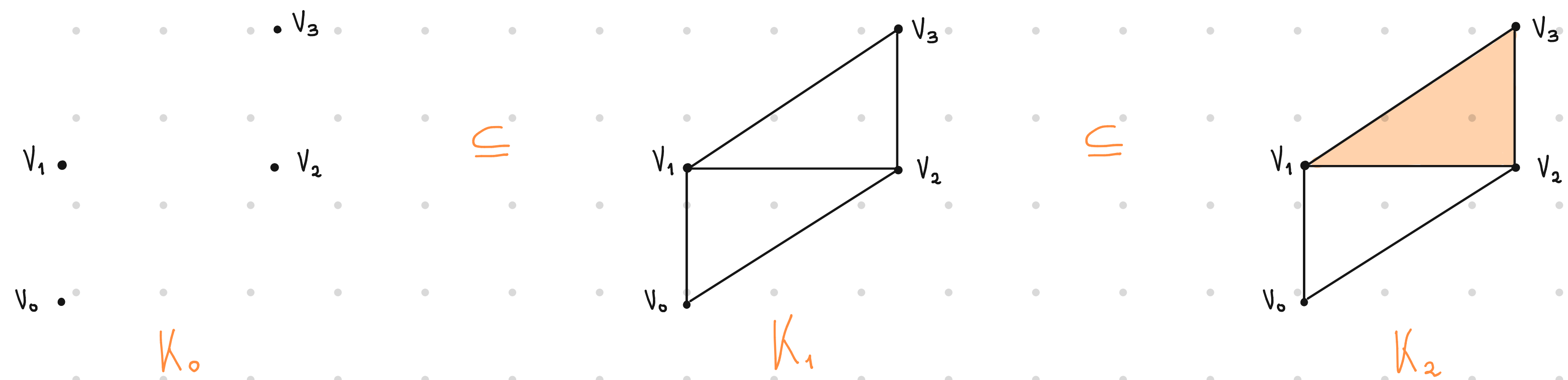
we'll always assume f_{ij} has finite rank

$$H_p^{i,j} := \operatorname{Im}\left(f_{ij}: H_p(K_i;\mathbb{F}_2) \longrightarrow H_p(K_j;\mathbb{F}_2)\right) \quad \text{for} \quad 0 \le i \le j \le r \quad (f_{ii} = \mathrm{id}).$$

② The **(i,j)-persistent p-th Betti number of K** is the dimension of $H_p^{i,j}$ over $\mathbb{F}_2$, i.e. $\beta_p^{i,j} := \dim_{\mathbb{F}_2}\left(H_p^{i,j}\right)$.

Remark: $$H_p^{i,j} = Z_p(K_i) \Big/ \big(B_{p+1}(K_j) \cap Z_p(K_i) \big) \qquad \forall\ 1 \leq i \leq j \leq r$$

Ex: Consider the following filtration and its p-th homology.



$\mathcal{F}$	K_0	$\hookrightarrow$	K_1	$\hookrightarrow$	K_2
$H_0(\mathcal{F})$	$\mathbb{F}_2^4$	$\xrightarrow[f_{01}]{[1\ 1\ 1\ 1]}$	$\mathbb{F}_2$	$\xrightarrow[f_{12}]{id}$	$\mathbb{F}_2$
$H_1(\mathcal{F})$	0		$\mathbb{F}_2^2$	$\xrightarrow[f_{12}]{[1\ 0]}$	$\mathbb{F}_2$
$H_p(\mathcal{F})$, $p \geq 2$	0		0		0

We can present $H_p^{i,j}$ and $\beta_p^{i,j}$ in a ij-plane for $0 \le i \le j \le 2$.

$H_p^{*,*}$

2	H_p^{02}	H_p^{12}	H_p^{22}
1	H_p^{01}	H_p^{11}	
0	H_p^{00}		
	0	1	2

$H_0^{*,*}$

2	$\mathbb{F}_2$	$\mathbb{F}_2$	$\mathbb{F}_2$
1	$\mathbb{F}_2$	$\mathbb{F}_2$	
0	$\mathbb{F}_2^4$		
	0	1	2

$\beta_0^{*,*}$

2	1	1	1
1	1	1	
0	4		
	0	1	2

$H_1^{*,*}$

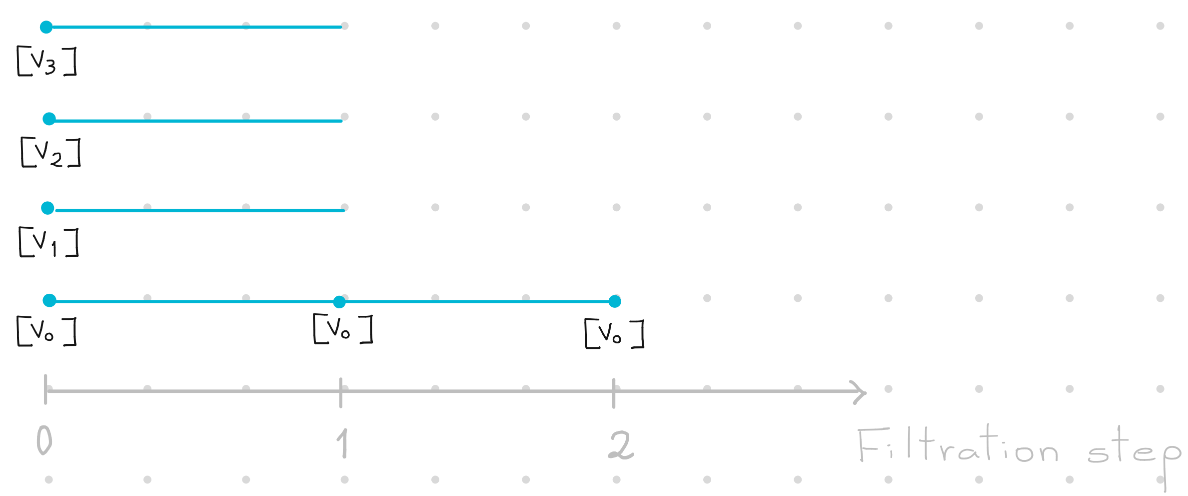
2	0	$\mathbb{F}_2$	$\mathbb{F}_2$
1	0	$\mathbb{F}_2^2$	
0	0		
	0	1	2

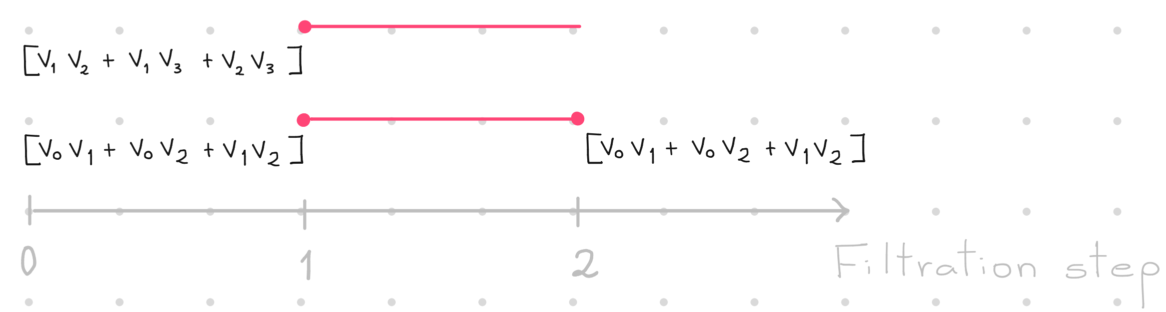
$\beta_1^{*,*}$

2	0	1	1
1	0	2	
0	0		
	0	1	2

The pth persistent homology of a filtered simplicial complex gives more refined information than just the homology of the single subcomplexes. We can visualize the information given by the vector spaces $H_p(K_i)$ together with the linear maps $f_{i,j}$ by drawing the following diagram: at filtration step i, we draw as many bullets as the dimension of the vector space $H_p(K_i)$. We then connect the bullets as follows: we draw an interval between bullet u at filtration step i and bullet v at filtration step $i+1$ if the generator of $H_p(K_i)$ that corresponds to u is sent to the generator of $H_p(K_{i+1})$ that corresponds to v. If the generator corresponding to a bullet u at filtration step i is sent to 0 by $f_{i,i+1}$, we draw an interval starting at u and ending at $i+1$.

Let's see an example for this filtration.

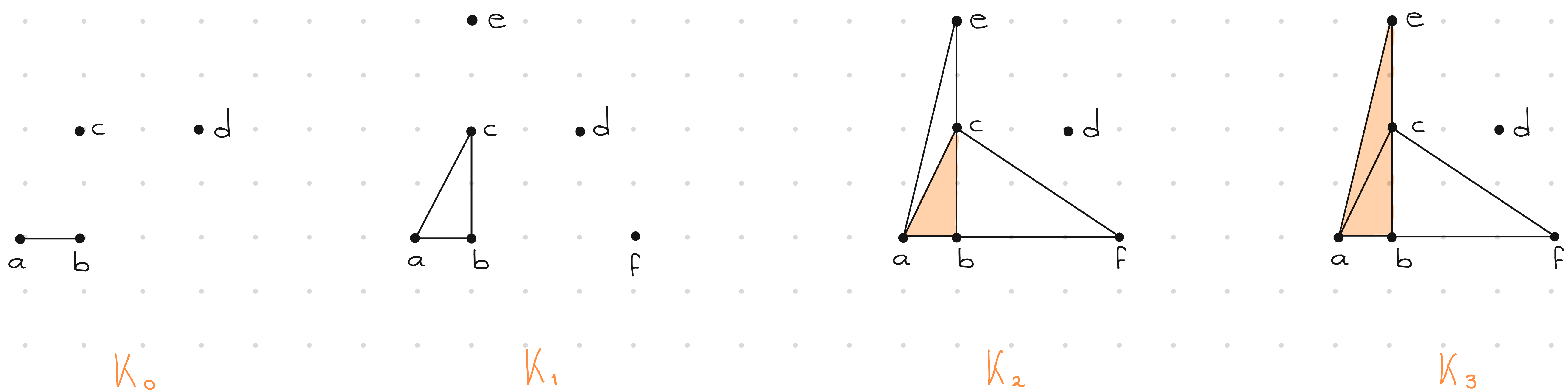




Such a diagram clearly depends on a choice of basis for the vector spaces $H_p(K_i)$, and a poor choice can lead to complicated and unreadable clutter. Fortunately, by the Fundamental Theorem of Persistent Homology there is a choice of basis vectors of $H_p(K_i)$ for each $i \in \{1, \ldots, l\}$ such that one can construct the diagram as a well-defined and unique collection of disjoint half-open intervals, collectively called a *barcode*.

→ see example below

Ex: Consider the following filtration



1ST CHOICE OF BASIS

Degree 0:

$$H_0(K_0) = \text{span}\{[a+c], [c], [d]\} \text{ where } [a]=[b]$$

$$H_0(K_1) = \text{span}\{[a], [a+f], [d], [e+f]\} \text{ where } [a]=[b]=[c]$$

$$H_0(K_2) = \text{span}\{[a], [d]\} \text{ where } [a]=[b]=[c]=[e]=[f]$$

$$H_0(K_3) = \text{span}\{[a], [d]\} \text{ where } [a]=[b]=[c]=[e]=[f]$$

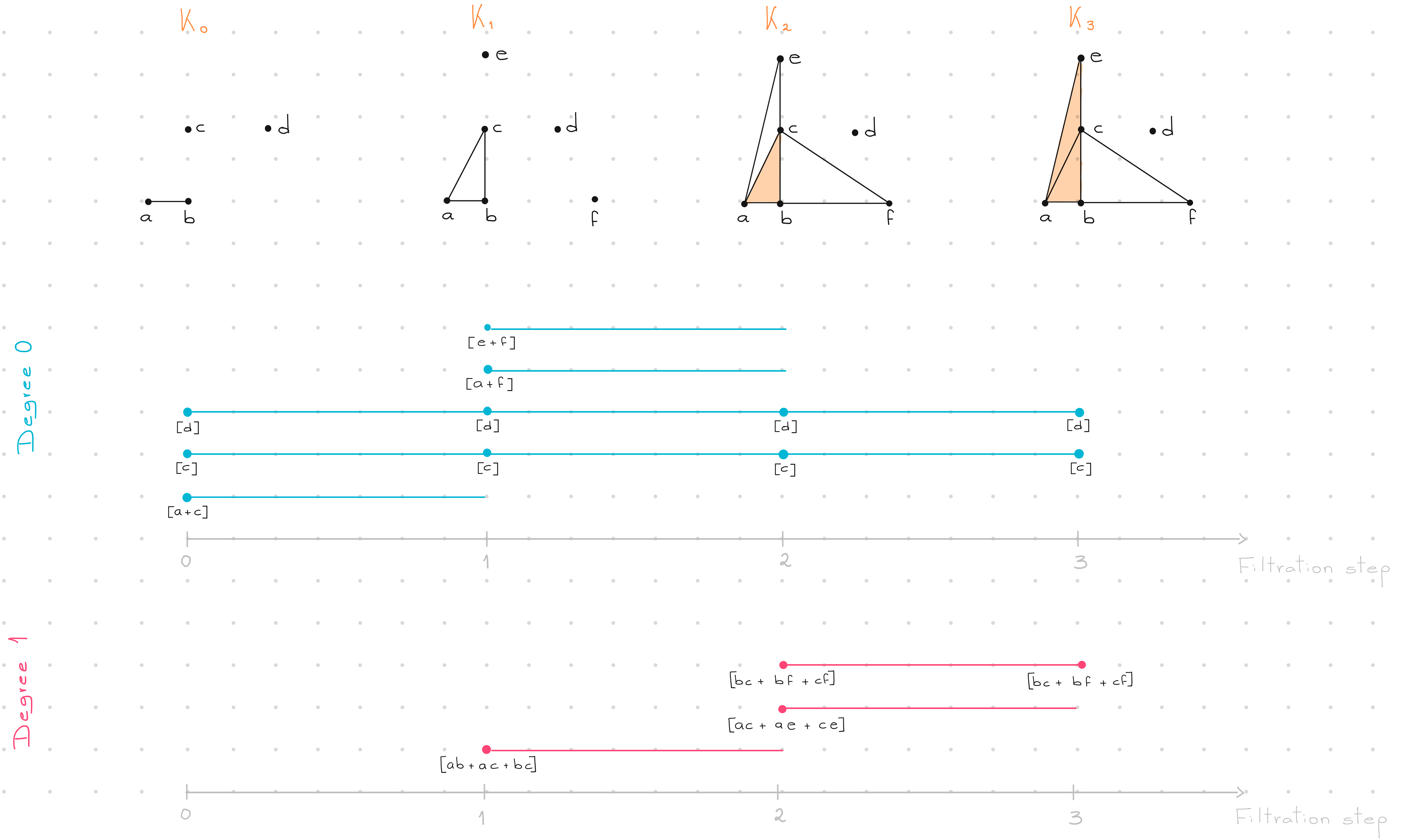
Degree 1:

$$H_1(K_0) = 0$$

$$H_1(K_1) = \text{span}\{[ab+ac+bc]\}$$

$$H_1(K_2) = \text{span}\{[ac+ae+ce], [bc+bf+cf]\}$$

$$H_1(K_3) = \text{span}\{[bc+bf+cf]\}$$



K0
K1
K2
K3
a
b
c
d
e
f
Degree 0
[e+f]
[a+f]
[d]
[c]
[a+c]
0
1
2
3
Filtration step
Degree 1
[bc + bf + cf]
[ac + ae + ce]
[ab + ac + bc]

2ND CHOICE OF BASIS

Degree 0:

$$H_0(K_0) = \text{span}\{[a], [c], [d]\} \quad \text{where} \quad [a] = [b]$$

$$H_0(K_1) = \text{span}\{[a], [d], [e], [f]\} \quad \text{where} \quad [a] = [b] = [c]$$

$$H_0(K_2) = \text{span}\{[a], [d]\} \quad \text{where} \quad [a] = [b] = [c] = [e] = [f]$$

$$H_0(K_3) = \text{span}\{[a], [d]\} \quad \text{where} \quad [a] = [b] = [c] = [e] = [f]$$